

# Spatial correlation of radar and gauge precipitation data in high temporal resolution

J. Brommundt and A. Bárdossy

Institute of Hydraulic Engineering, Chair of Hydrology and Geohydrology, Universitaet Stuttgart, Germany

Received: 17 August 2006 – Revised: 9 January 2007 – Accepted: 5 March 2007 – Published: 26 April 2007

**Abstract.** A multi-sites precipitation time series generator for engineering designs is currently being developed. The objective is to generate several time series' simultaneously with correct inter-station relationships. Therefore, a model to estimate correlation between stations for arbitrary points in a project area is needed, using rain gauge data as well as radar data.

Two methods are applied to compare the spatial behaviour of precipitation in both the rain gauge data and the radar data. The first approach is to calculate precipitation intensities from radar reflectivity and use it as gauge data. The results show that the spatial structure in both data sets is similar, but cross correlation varies too much to use radar derived spatial correlation to describe gauge inter-station relationship. Thus, a second approach was tested to account for the differences in the spatial correlation associated to the distribution. Using the indicator time series, cross correlations for different quantiles were calculated from both the rain gauge and radar data. This approach shows that cross correlation varies depending on the chosen quantile. In the lower quantiles, the correlation is very similar in rain gauge and radar data, hence a transfer is possible. This insight is useful to derive cross correlations of rain gauges from radar images. Correlation data for rain gauges thus obtained contains all the information about heterogeneity and anisotropy of the spatial structure of rainfall, which is in the radar data.

## 1 Introduction

The state of Baden-Wuerttemberg (BW) in southern Germany has a diverse topography ranging from 84 m a.s.l. to 1 405 m a.s.l. with an average elevation of 502 m a.s.l.. The dominating precipitation process is advective rain cells arriv-

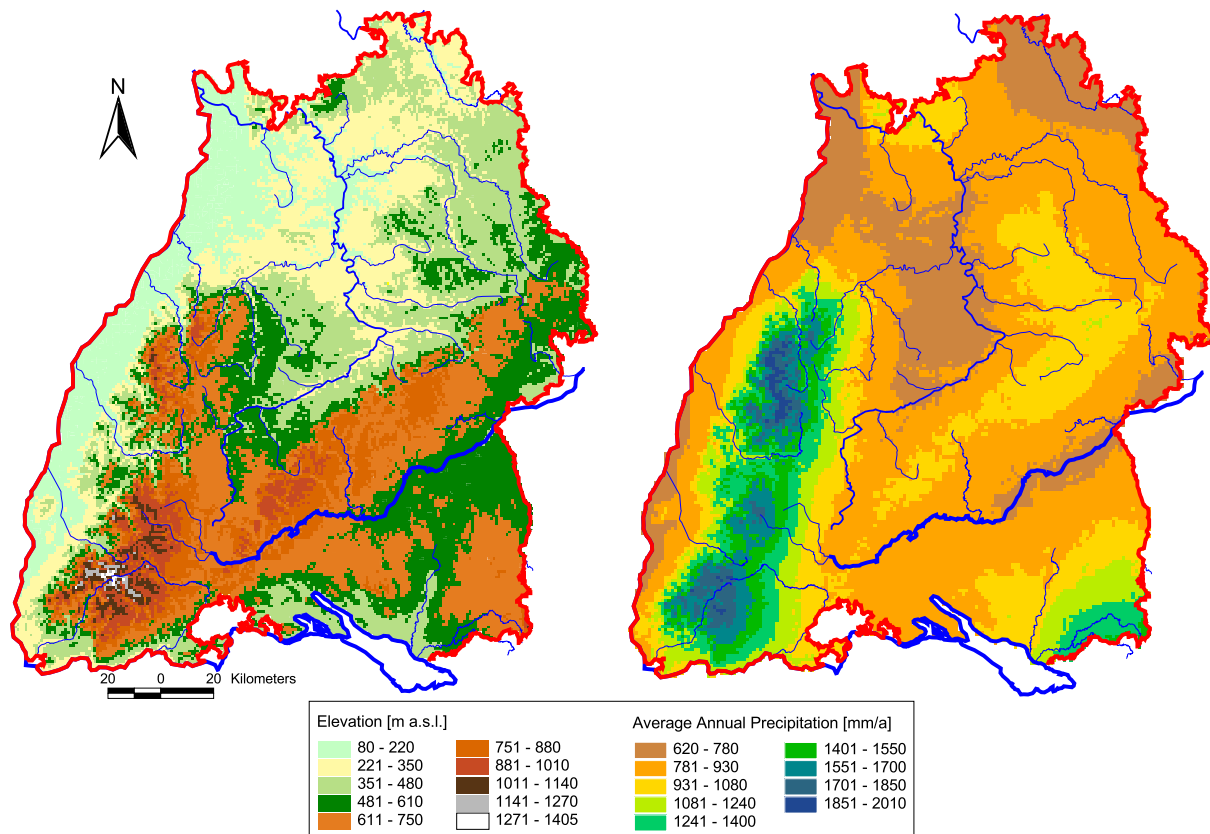
ing from the south-west. Some of the highest mountains of BW are located in the Black Forest, which stretches along the Rhine valley from south to north. As such, they receive the majority of the incoming rain. Because of topography and the predominant weather direction precipitation is very heterogeneous in BW, as shown in Fig. 1.

In the layout of hydraulic structures, such as sewage or urban drainage systems, hydraulic simulation models are state-of-the-art. These models require complete precipitation time series' as input, which should be about 30 years long and have typically a temporal resolution of 5 min. Unfortunately, there is not a sufficient number of rain gauge time series', to fulfil these demands, especially if spatial variability is considered. To account for this discrepancy, a stochastic precipitation time series generator has been developed and incorporated into a GIS scheme. This time series generator implements a generation method suggested by Bárdossy (1998) and covers the whole territory of BW, which has a total area of 35 751 km<sup>2</sup>. The point precipitation time series generator incorporates local statistical properties, which are derived from measured rain data at gauge sites and then interpolated on a 1 × 1 km<sup>2</sup> grid using External Drift Kriging (Ahmed and de Marsily, 1987).

Experience with these time series' has shown that a single point is not representative for a sewage system that drains a larger area. Therefore, the objective is to expand the generator's capabilities to generate several simultaneous time series'. To accommodate this, the spatial relationship of rain at gauges needs to be estimated for arbitrarily chosen point pairs (stations) in the project region, i.e. a regionalisation model for the spatial relationship between station pairs must be established. For daily data, Brommundt and Bárdossy (2005) have presented a regionalisation model based on four-dimensional kriging. Additionally, a model for the inter-station spatial relationship of hourly rain data is needed, which is a necessary input for the generator.

---

Correspondence to: J. Brommundt  
(juergen.brommundt@iws.uni-stuttgart.de)



**Fig. 1.** Elevation (*left*) and annual average precipitation (*right*) in Baden-Wuerttemberg.

The spatial relationship of rainfall data is an input for many hydrological and meteorological applications. Whenever spatial rainfall is simulated, an underlying structure must be assumed. In many cases, correlation or covariance are used to characterise or measure spatial association (Pegram and Clothier, 2001; Favre et al., 2002; Koutsoyiannis et al., 2003). The inter-station correlation is neither homogeneous nor isotropic (Zawadzki, 1973), but depends on the location and on the orientation of each station pair (Lanza et al., 2001). Sampson (1986) handles this problem in a model through deformations of the geographic coordinate system. The aim is to find a geographic plane in which the correlation can be assumed isotropic. Correlation can be estimated by a simple functional approach within that plane. Monestiez et al. (1993) applied this technique using radial basis deformations, while Meiring et al. (1997) used splines to find the appropriate deformed plane. All of these methods are limited to small areas.

Weather radars allow rainfall in a high spatial and temporal resolution to be measured (Collier, 1996). Krajewski et al. (1996) show that the radar observation process has an influence on the different statistics of rainfall. Regarding correlation formulated by the decorrelation length, the influence of the radar rainfall estimation algorithm is rather

small, whereas temporal aggregation has a larger influence. They compared the correlation derived from radar data to correlation from two rainfall models and find that radar is more smoothing and overestimates correlation of the modelled "true" rainfall.

Gebremichael et al. (2004) compared the capability of radar to catch the small-scale variability of rain with respect to correlation. They used a network of rain gauges as reference and concluded that radar estimates of correlations are prone to errors up to 30%. This study assumed isotropic correlation. Moszkowicz (2000) performed a similar study finding strong anisotropy in the correlation structure of rainfall. He uses rain gauge and radar data to describe correlation from short range up to several tens of kilometers.

These findings lead to the conclusion that radar images are a good way to gain information about variability and correlation of rainfall. Anisotropy and heterogeneity are well reflected, but statistics themselves have to be handled with special care. Therefore, we try to use radar data to estimate cross correlation of rain gauges. This way heterogeneity and anisotropy of the cross correlation do not need to be modelled, but can be taken from the radar data, which is available in a high spatial resolution.

**Table 1.** Transformations of reflectivity into rain intensity.

Cl.	Range [dBz]	A [mm/h]	B [mm/h]	C [mm/h]
0	0–7	0.000	0.000	0.000
1	7–19	0.181	0.251	0.200
2	19–28	0.759	1.163	1.500
3	28–37	2.944	5.005	3.900
4	37–46	11.419	21.538	14.000
5	46–55	44.293	92.685	50.000
6	55–60	109.991	150.000	100.000

(A) Z-R-relationship,  $a=473.327$ ,  $b=1.529$  (own fit)

(B) Z-R-relationship,  $a=256$ ,  $b=1.42$  (DWD standard)

(C) Given class wise transformation, (Wetter.com, 2006)

## 2 Inter-station spatial relationship

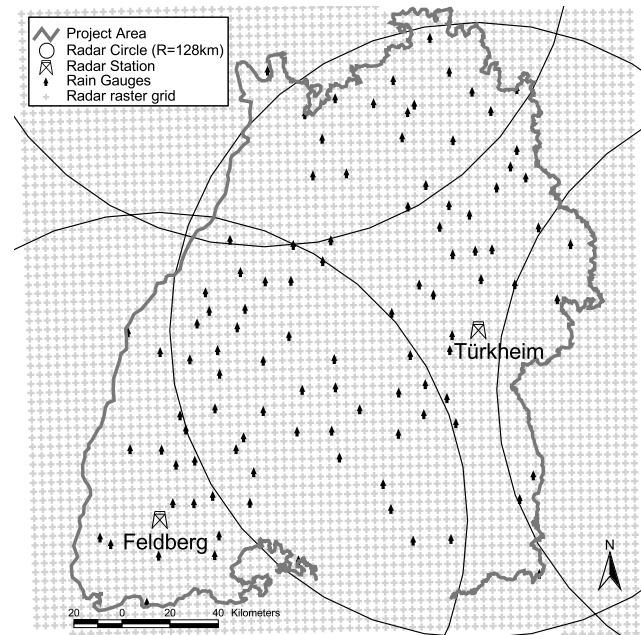
In the generation scheme, we formulate the spatial relationship between stations by the cross correlation

$$r(x, y) = \frac{\sum_T (h_x(t) - \bar{h}_x) (h_y(t) - \bar{h}_y)}{\sqrt{\sum_T (h_x(t) - \bar{h}_x)^2} \sqrt{\sum_T (h_y(t) - \bar{h}_y)^2}}. \quad (1)$$

Equation (1) describes the correlation between the gauges at locations  $x$  and  $y$  calculated from the time series  $h_i(t)$  with the time series length  $T$ . The basic assumption for the application of cross correlation as a measure for spatial relationship is the normality of the input data. In general, this is not the case for rainfall data (Habib et al., 2001). Nevertheless, correlation has been used frequently. Zawadzki recommends using logarithmic transformed precipitation (Kessler and Neas, 1994) to reduce the skewness of the rainfall data. This transformation can only be applied to the wet part of a rainfall time series. Since we want to describe the relationship of the whole time series, we use traditional cross correlation, keeping in mind that special attention has to be given to this point.

To estimate the correlation between rain data measured at different gauges the data needs to be registered synchronous. The higher the temporal resolution of gauge data, the greater the importance of synchrony. In our dataset (Sect. 2.1), synchrony cannot be guaranteed since every gauge has its own time keeper. This is most likely the reason for the inapplicability of the kriging approach presented in Brommundt and Bárdossy (2005) to the interpolation of cross correlation of hourly data. Application was only tested with the rain gauge data. Tests with higher and lower temporal aggregations failed as well.

Radar images cover this gap, as they are considered synchronous images of rain. Another advantage is the spatial information in a high resolution, so there is no need for a post processing like a regionalisation or an interpolation scheme.

**Fig. 2.** Map of the rain gauges and radar stations within the project area.

However, radar observations are in a plane with increasing altitude above ground the further the beam gets from the station. Furthermore, it provides a spatially aggregated image of the precipitation over an area typically in the order of magnitude of square kilometers. Rain gauges are point measurements based upon a funnel of some hundred square centimeters catch area (Collier, 1996). Using radar for rainfall estimation must consider these discrepancies (Gjertsen et al., 2003).

### 2.1 Available data

German Weather Service (DWD) runs a network of 16 C-band Doppler radars covering Germany (DWD, 2003). They provide several composite radar images in different temporal and spatial resolutions. For our purposes, long time series' in combination with a temporal resolution less than or equal to 1 h are important. Therefore, DWD provided the national composite image PC (= Picture Composite) in 15 min resolution on a  $4 \times 4 \text{ km}^2$  grid from 1 July 1997 until 31 December 2004. These images contain censored reflectivity data [dBz] in seven classes (Table 1). The reflectivity shown is measured in the lowest beam elevation between  $0.5^\circ$  and  $1.8^\circ$ , which varies according to local topography. With an overall failure rate of 1.55% the data quality is very good.

Out of the rain gauge data set available from the setup of the simulation scheme, 101 gauging stations in a 5 min resolution were chosen which have a good data quality in the time span of the radar images. They are ombrometer stations using the weighing principle, also operated by DWD.

**Table 2.** Average frequencies of the radar and rain gauge data in the different classes.

data	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
15 min radar	8.38E-1	1.03E-1	4.37E-2	1.31E-2	2.03E-3	3.76E-4	4.2E-5
15 min rain gauges	8.86E-1						

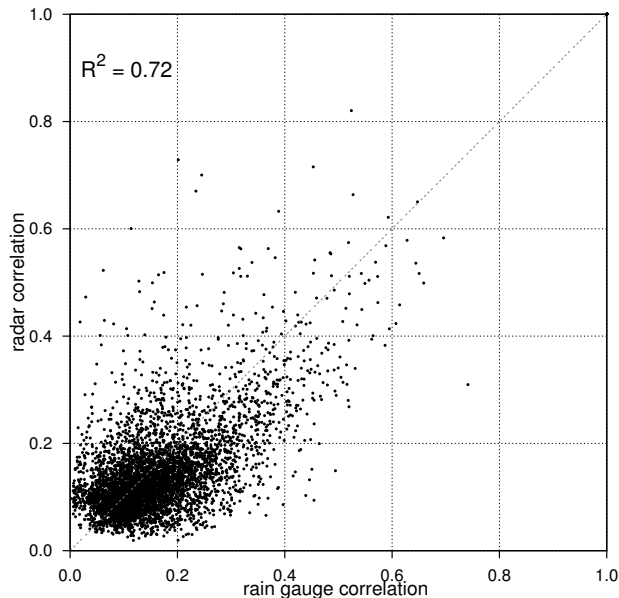
**Fig. 3.** Scatterplot of rain and radar derived cross correlation.

Figure 2 shows the project area with the radar raster grid, the rain gauges and the radar stations. The spatial coverage of the rain gauges is quite sufficient and homogeneous. Four radar stations cover the project area as marked by the radar circles. Two of them, Feldberg and Tuerkheim, are centered within the area. Since we use composite radar images, the intersections in which two radar beams overlap each other are of particular importance. DWD generates composite images using the highest value from overlapping radar images for raster points in the intersection. Because of this, the radius between radar station and raster point in intersections cannot be determined, because it is unknown which radar station's information has been used. This lack of information limits evaluation possibilities.

## 2.2 Straightforward approach

The most obvious and easiest way to gain correlation from radar images for rain gauges is to estimate rain intensities from the radar reflectivities. A common way to do this is to fit a  $Z$ - $R$ -relationship (Collier, 1996), which links the rain

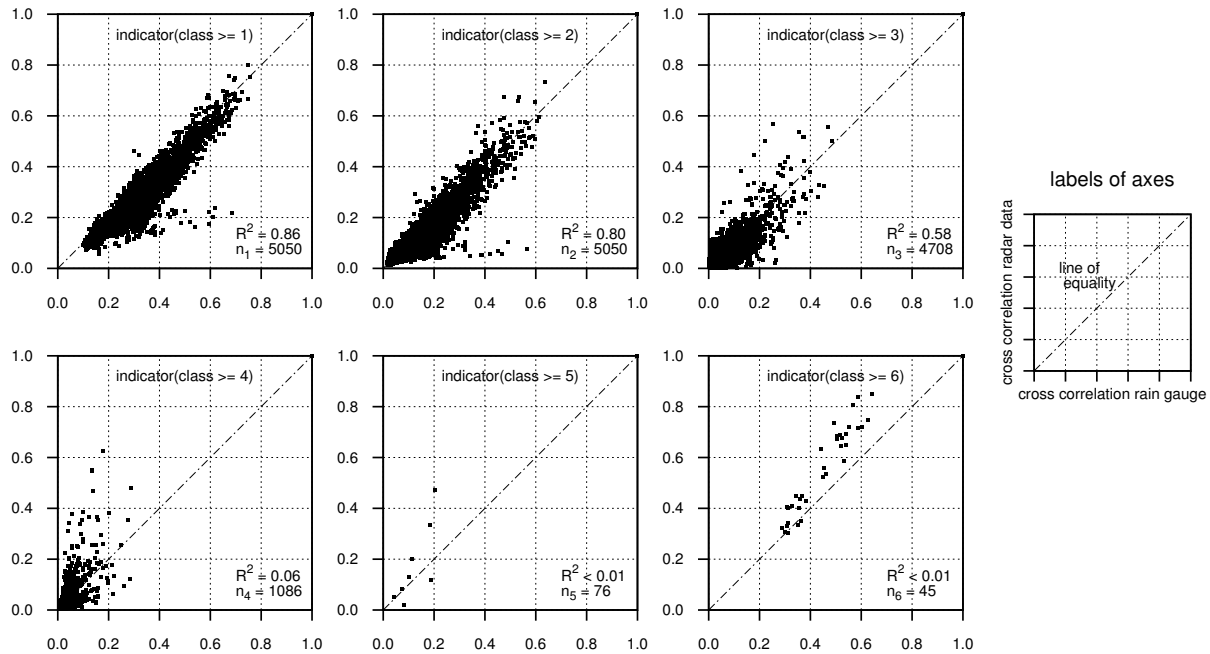
intensity  $R$  to the radar reflectivity  $Z$ :

$$R = \left( \frac{Z}{a} \right)^{\frac{1}{b}} = \left( \frac{10^{\frac{\text{dBZ}}{10}}}{a} \right)^{\frac{1}{b}}. \quad (2)$$

Another advantage of this formulation is that rain intensities can be better processed, i.e., they can be aggregated. From the data set, the parameters  $a$  and  $b$  of Eq. (2) can be estimated, and was done at the 101 gauging stations. Since the raw reflectivity data of the radar cannot be aggregated, we used 15 min data from radar and rain gauges. At this point, the closest radar raster point was associated to each gauge. Trials with averages from 4 to 9 closest neighbours gained no improvements. To follow Zawadzki (1975), the average over 9 neighbours should be used for data in 15 min time resolution assuming a speed of storm motion of 1 km/min. Since results did not differ, we used only the closest neighbour.

Due to the different fractions of dry times, i.e.  $R=0$  mm/h (Table 2),  $a$  and  $b$  were fitted according to the distribution in radar and rain gauged data: The class boundaries of the radar data are known as reflectivities. The rain gauge data has to be filtered in such a way that the distribution of rain and radar are as close as possible. One has to consider that the lowest class must have an error, so the censoring should start with the highest values. Table 2 shows that the average frequency of Class 0 in the radar data is 83.8%, while the rain gauges register in 88.6% of the time with no rain. With this, the associated boundaries of the rain gauge and the radar data classes can be used to derive  $a$  and  $b$  in an optimisation scheme. Assuming the power law given in Eq. (2) and the known values of reflectivity  $Z$  (Table 1) at the boundaries,  $a=473.327$  and  $b=1.529$  were gained. From  $a$  and  $b$ , the boundaries of the rain intensities of the classes in the rain gauge data can be calculated. Class wise transformation is given in Table 1 column (A).

To evaluate the established relationship, two standard methods were applied simultaneously. DWD supplies  $a=256$  and  $b=1.42$  as standard values for Eq. (2) (Bartels et al., 2004), which are evaluated in column (B) of Table 1. The supplied radar images are also published (Wetter.com, 2006). From this source, a class wise transformation is provided, which can be found in column (C) of Table 1.



**Fig. 4.** Class wise cross correlation for the indicator correlation.

All three transformations were used to calculate the cross correlation (Eq. 1) of any station pair in the data set. Since we want to know correlation of hourly data, all analyses were carried out with the hourly aggregated data.

Figure 3 shows the results for fitted transformation (A). The coefficient of determination between cross correlations from rain gauges and radar images is  $R^2=0.72$ , but radar seems to underestimate the correlation estimated from the rain data systematically. Transformations (B) and (C) lead to similar results. Further problems are an overestimation of rain totals in the radar data and a lack of synchrony within the rain gauge data and between rain gauge and radar data.

In the straightforward approach, the differences in the distributions of rain gauge and radar data are neglected. This may be one of the reasons that the correlation is not transferable in a straightforward manner. The non-normality of the data is an additional shortcoming of this approach. Another reason may be that cross correlation measures the degree of association over the whole distribution, whereas the relationship might vary over the distribution. Such differences are smoothed out in the cross correlation presented in Fig. 3. Therefore, a different approach was used, which allows to account for the differences in the distributions and avoids the problem of non-normality.

### 2.3 Cross correlation of quantiles

When the cross correlation of the distribution for different quantiles is needed, a common approach is to calculate the indicator time series from the original time series. Every

value  $c(t)$  of the time series, which is higher than a selected reference value  $c_{\text{ref}}$  is assigned the value 1 in the indicator time series  $I(t)$ , otherwise the value is 0. Equation (3) shows the processing for the data set, from which we used classes 1 to 6.

$$I(t, c_{\text{ref}}) = \begin{cases} 1 & \text{if } c(t) \geq c_{\text{ref}} \\ 0 & \text{else} \end{cases}, c_{\text{ref}} = 1, \dots, 6. \quad (3)$$

Now the cross correlation (Eq. 1) of  $I(t)$  is calculated. The cross correlation of the indicator series can be interpreted as the cross correlation of the quantile  $P(c > c_{\text{ref}})$ . If the cross correlation for different reference values is evaluated, we gain the correlation for different quantiles. Class 0 is neglected in this procedure since the indicator would be always 1 and correlation cannot be calculated. The reference value  $c_{\text{ref}}$  for each quantile was derived from the class boundaries in Table 1. Since the main interest is the relationship of hourly values, we aggregated the original data to hourly sums. In the aggregated time series, these reference values have to be adapted to keep the density of the non-aggregated data. Using the quantiles of radar and rain distribution, the reference values  $c_{\text{ref}}$  for the rain gauge data were calculated. Thus, the same quantiles were used for indicator cross correlation estimation of rain gauge and radar data, which allows for a class-wise intercomparison.

Figure 4 shows the cross correlations for all six classes as scatter plots. On the x-axis, the cross correlation calculated from the rain gauge indicator time series' is plotted, while the y-axis shows the correlation derived from the radar indicator series'. The different plots show that cross correlation

varies for different quantiles. The higher values are not so frequently observed. The number of 1s in the indicator series is smaller, but calculation of the cross correlation is only meaningful if at least some events are measured. Hence, the number of points in the scatter plots in Fig. 4 decreases with increasing class number (Table 2). The relationship is strong in the classes one and two. The coefficient of determination is much higher than in the straightforward approach (Sect. 2.2, Fig. 3). In the higher classes, there is still a relationship, but it is weaker than in the lower classes.

### 3 Conclusions and outlook

Two approaches were shown to link inter-station relationships derived from rain gauge data to relationships calculated from radar data. First, a straightforward approach applying a  $Z$ - $R$ -transformation from radar reflectivity to rain demonstrated that the spatial structure of precipitation is similar in both datasets. The second more sophisticated approach assumes that the cross correlation depends on the distribution. The results show that this assumption is correct. Unfortunately, the data set quality does not allow for a quantification of this correlation.

In any case, the correlation of rain gauge and radar indicator time series' are very similar in the lower classes, indicating that the pattern of the inter-station correlation is also similar in radar and rain gauges. Hence, radar images as spatial information can supplement rain gauge data, which are always point measurements. Regionalisation becomes dispensable since radar images cover the whole area of interest. The next step is to link the cross correlation of the quantiles back to the cross correlation of the whole time series.

This study used composite radar images since coverage of the whole project area was needed. Additionally these images were the only data available for a long time (7.5 years). Therefore, the poor resolution in reflectivity with only seven classes was accepted. To verify the results, this survey should be repeated with radar data from one single radar station in higher reflectivity resolution. A quantitative estimation of the variation of the cross correlation with the quantile can therefore be derived from that survey. The agreement in the correlation between rain gauge and radar data is expected to increase.

**Acknowledgements.** This work is funded by the "Baden-Wuerttemberg Research Program Securing a Sustainable Living Environment" (BWPLUS) at the Forschungszentrum Karlsruhe with funds of the State of Baden-Wuerttemberg. The authors thank the German Weather Service (DWD) and the State Institute for Environmental Protection Baden-Wuerttemberg (LUBW) for data supply and cooperation. The useful comments of two reviewers were valuable advice.

Edited by: S. C. Michaelides and E. Amitai

Reviewed by: anonymous referees

### References

- Ahmed, S. and de Marsily, G.: Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, *Water Resour. Res.*, 23(9), 1717–1737, 1987.
- Bárdossy, A.: Generating precipitation time series using simulated annealing, *Water Resour. Res.*, 34, 1737–1744, 1998.
- Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N.: Projekt RADOLAN, Routineverfahren zur Online-Aneicherung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Abschlussbericht (in German), Deutscher Wetterdienst, 2004.
- Brommundt, J. and Bárdossy, A.: Stochastic Generation of Synthetic Precipitation Time Series with High Temporal and Spatial Resolution for Engineering Practice, in: *Geophysical Research Abstracts*, EGU General Assembly, Vol. 7, [http://www.iws.uni-stuttgart.de/institut/hydrologie/publikationen/bro\\_bard\\_egu2005.pdf](http://www.iws.uni-stuttgart.de/institut/hydrologie/publikationen/bro_bard_egu2005.pdf), 2005.
- Collier, C. G.: Applications of weather radar systems: a guide to the use of radar data in meteorology and hydrology, Wiley-Praxis series in atmospheric physics, Chichester, 2 edn., 1996.
- DWD: Der Radarverbund des Deutschen Wetterdienstes, Tech. rep., Deutscher Wetterdienst, <http://www.dwd.de/de/Technik/Datengewinnung/Radarverbund/Radarbroschuere.pdf>, 2003.
- Favre, A.-C., Musy, A., and Morgenthaler, S.: Two-site modeling of rainfall based on the Neyman-Scott process, *Water Resour. Res.*, 38, 1–43, doi:10.1029/2002WR001343, 2002.
- Gebremichael, M., Krajewski, W., and Ciach, G.: Assessment of the statistical characterization of small-scale rainfall variability from radar, in: *Proceedings of the Sixth International Symposium on Hydrological Applications of Weather Radar*, 2004.
- Gjertsen, U., Salek, M., and Michelson, D. B.: Gauge-adjustment of radar-based precipitation estimates – a review, in: *COST-717 working document No. WDD 02 200310 1*, <http://www.smhi.se/cost717/>, 2003.
- Habib, E., Krajewski, W., and Ciach, G.: Estimation of Rainfall Interstation Correlation, *J. Hydrometeorol.*, 2, 621–629, doi:10.1175/1525-7541(2001)002<0621:EORIC>2.0.CO;2, 2001.
- Kessler, E. and Neas, B.: On correlation, with applications to the radar and raingage measurement of rainfall, *Atmos. Res.*, 34, 217–229, 1994.
- Koutsoyiannis, D., Onof, C., and Wheeler, H. S.: Multivariate rainfall disaggregation at a fine timescale, *Water Resour. Res.*, 39, 1173, doi:10.1029/2002WR001600, 2003.
- Krajewski, W., Anagnostou, E., and Ciach, G.: Effects of the radar observation process on inferred rainfall statistics, *J. Geophys. Res.*, 101, 26 493–26 502, 1996.
- Lanza, L. G., Ramírez, J. A., and Todini, E.: Stochastic rainfall interpolation and downscaling, *Hydrol. Earth Syst. Sci.*, 5, 139–143, 2001, <http://www.hydrol-earth-syst-sci.net/5/139/2001/>.
- Meiring, W., Monestiez, P., Sampson, P., and Guttorp, P.: Developments in the modelling of nonstationary spatial covariance structure from space-time monitoring data, in: *Geostatistics Wollongong '96*, edited by: Baafi, E. and Schofield, N., Vol. 1, pp. 162–173, Kluwer Academic Publishers, 1997.
- Monestiez, P., Sampson, P. D., and Guttorp, P.: Modelling of heterogeneous spatial correlation structure by spatial deformation, in: *Cahiers de Geostatistique*, École Nationale Supérieure des Mines de Paris, 1993.

- Moszkowicz, S.: Small-Scale Structure of Rain Field-Preliminary Results Basing on a Digital Gauge Network and on MRL-5 Legionowo Radar, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans Atmos.*, 25, 933–938, 2000.
- Pegram, G. G. S. and Clothier, A. N.: Downscaling rainfields in space and time, using the String of Beads model in time series mode, *Hydrol. Earth Syst. Sci.*, 2, 175–186, 2001, <http://www.hydrol-earth-syst-sci.net/2/175/2001/>.
- Sampson, P. D.: Spatial covariance estimation by scaled-metric scaling and biorthogonal grids, Tech. Rep. Technical Report No. 91, Department of Statistics, University of Washington, 1986.
- Wetter.com: Niederschlags-Radar Deutschland, Online, <http://www.wetter.com/v2/?SID=&LANG=DE&LOC=0270&LOCFROM=0270>, 2006.
- Zawadzki, I.: On radar-raingage comparison, *J. Appl. Meteorol.*, 14, 1430–1436, 1975.
- Zawadzki, I. I.: Statistical properties of precipitation patterns, *J. Appl. Meteorol.*, 12, 459–472, doi:10.1175/1520-0450(1973)012<0459:SPOPP>2.0.CO;2, 1973.